

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Lal, Preeti
Hillman, Jennifer L.
Guegler, Karl J.
Corley, Neil C.
Baughn, Mariah R.
Azimzai, Yalda

(ii) TITLE OF THE INVENTION: HUMAN UBIQUITIN-CONJUGATING ENZYME
HOMOLOGS

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: TO BE ASSIGNED
(B) FILING DATE: HERewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Cerrone, Michael C
(B) REGISTRATION NUMBER: 39,132
(C) REFERENCE/DOCKET NUMBER: PF-0502 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650-855-0555
(B) TELEFAX: 650-845-4166
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 250 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: PROSNOT14
(B) CLONE: 1728211

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1 :

Met Ala Leu Leu Ala Thr Ser Leu Pro Glu Gly Ile Met Val Lys
5 10 15

09930026-081401

PF-0502-1 DIV

Thr	Phe	Glu	Asp	Arg	Met	Asp	Leu	Phe	Ser	Ala	Leu	Ile	Lys	Gly	
				20					25					30	
Pro	Thr	Arg	Thr	Pro	Tyr	Glu	Asp	Gly	Leu	Tyr	Leu	Phe	Asp	Ile	
				35					40					45	
Gln	Leu	Pro	Asn	Ile	Tyr	Pro	Ala	Val	Pro	Pro	His	Phe	Cys	Tyr	
				50					55					60	
Leu	Ser	Gln	Cys	Ser	Gly	Arg	Leu	Asn	Pro	Asn	Leu	Tyr	Asp	Asn	
				65					70					75	
Gly	Lys	Val	Cys	Val	Ser	Leu	Leu	Gly	Thr	Trp	Ile	Gly	Lys	Gly	
				80					85					90	
Thr	Glu	Arg	Trp	Thr	Ser	Lys	Ser	Ser	Leu	Leu	Gln	Val	Leu	Ile	
				95					100					105	
Ser	Ile	Gln	Gly	Leu	Ile	Leu	Val	Asn	Glu	Pro	Tyr	Tyr	Asn	Glu	
				110					115					120	
Ala	Gly	Phe	Asp	Ser	Asp	Arg	Gly	Leu	Gln	Glu	Gly	Tyr	Glu	Asn	
				125					130					135	
Ser	Arg	Cys	Tyr	Asn	Glu	Met	Ala	Leu	Ile	Arg	Val	Val	Gln	Ser	
				140					145					150	
Met	Thr	Gln	Leu	Val	Arg	Arg	Pro	Pro	Glu	Val	Phe	Glu	Gln	Glu	
				155					160					165	
Ile	Arg	Gln	His	Phe	Ser	Thr	Gly	Gly	Trp	Arg	Leu	Val	Asn	Arg	
				170					175					180	
Ile	Glu	Ser	Trp	Leu	Glu	Thr	His	Ala	Leu	Leu	Glu	Lys	Ala	Gln	
				185					190					195	
Ala	Leu	Pro	Asn	Gly	Val	Pro	Lys	Ala	Ser	Ser	Ser	Pro	Glu	Pro	
				200					205					210	
Pro	Ala	Val	Ala	Glu	Leu	Ser	Asp	Ser	Gly	Gln	Gln	Glu	Pro	Glu	
				215					220					225	
Asp	Gly	Gly	Pro	Ala	Pro	Gly	Glu	Ala	Ser	Gln	Gly	Ser	Asp	Ser	
				230					235					240	
Glu	Gly	Gly	Ala	Gln	Gly	Leu	Ala	Phe	Ser						
				245					250						

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 282 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: SINTNOT13
- (B) CLONE: 1803905

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2 :

Met	Ala	Gln	Gln	Gln	Met	Thr	Ser	Ser	Gln	Lys	Ala	Leu	Met	Leu	
				5					10					15	
Glu	Leu	Lys	Ser	Leu	Gln	Glu	Glu	Pro	Val	Glu	Gly	Phe	Arg	Ile	
				20					25					30	
Thr	Leu	Val	Asp	Glu	Ser	Asp	Leu	Tyr	Asn	Trp	Glu	Val	Ala	Ile	
				35					40					45	
Phe	Gly	Leu	Pro	Asn	Thr	Leu	Tyr	Glu	Gly	Gly	Tyr	Phe	Lys	Ala	
				50					55					60	
His	Ile	Lys	Phe	Pro	Ile	Asp	Tyr	Pro	Tyr	Ser	Pro	Pro	Thr	Phe	
				65					70					75	
Arg	Phe	Leu	Thr	Lys	Met	Trp	His	Pro	Asn	Ile	Tyr	Glu	Asn	Gly	
				80					85					90	
Asp	Val	Cys	Ile	Ser	Ile	Leu	His	Pro	Pro	Val	Asp	Asp	Pro	Gln	
				95					100					105	
Ser	Gly	Glu	Leu	Pro	Ser	Glu	Arg	Trp	Asn	Pro	Thr	Gln	Asn	Val	
				110					115					120	

09930026-081401

Arg	Thr	Ile	Leu	Leu	Ser	Val	Ile	Ser	Leu	Leu	Asn	Glu	Pro	Asn	
				125					130					135	
Thr	Phe	Ser	Pro	Ala	Asn	Val	Asp	Ala	Ser	Val	Met	Phe	Arg	Lys	
				140					145					150	
Trp	Arg	Asp	Ser	Lys	Gly	Lys	Asp	Lys	Glu	Tyr	Ala	Glu	Ile	Ile	
				155					160					165	
Arg	Lys	Gln	Val	Ser	Ala	Thr	Lys	Ala	Glu	Ala	Glu	Lys	Asp	Gly	
				170					175					180	
Val	Lys	Val	Pro	Thr	Thr	Leu	Ala	Glu	Tyr	Cys	Ile	Lys	Thr	Lys	
				185					190					195	
Val	Pro	Ser	Asn	Asp	Asn	Ser	Ser	Asp	Leu	Leu	Tyr	Asp	Asp	Leu	
				200					205					210	
Tyr	Asp	Asp	Asp	Ile	Asp	Asp	Glu	Asp	Glu	Glu	Glu	Glu	Asp	Ala	
				215					220					225	
Asp	Cys	Tyr	Asp	Asp	Asp	Asp	Ser	Gly	Met	Arg	Ser	Arg	Asp	Val	
				230					235					240	
Leu	Leu	Gln	Cys	Pro	Cys	Thr	Ala	Leu	Pro	Ser	Gln	Ala	Lys	Gly	
				245					250					255	
Arg	Gly	Ala	Ser	Gly	Asp	Leu	Ala	Met	Ala	Pro	Gln	Gln	Lys	Pro	
				260					265					270	
Ile	His	Ser	Gly	Trp	Gly	Asn	Thr	His	Ser	Ser	Cys				
				275					280						

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 318 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: COLNTUT16
- (B) CLONE: 2792472

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3 :

Met	Glu	Thr	Arg	Tyr	Asn	Leu	Lys	Ser	Pro	Ala	Val	Lys	Arg	Leu	
				5					10					15	
Met	Lys	Glu	Ala	Ala	Glu	Leu	Lys	Asp	Pro	Thr	Asp	His	Tyr	His	
				20					25					30	
Ala	Gln	Pro	Leu	Glu	Asp	Asn	Leu	Phe	Glu	Trp	His	Phe	Thr	Val	
				35					40					45	
Arg	Gly	Pro	Pro	Asp	Ser	Asp	Phe	Asp	Gly	Gly	Val	Tyr	His	Gly	
				50					55					60	
Arg	Ile	Val	Leu	Pro	Pro	Glu	Tyr	Pro	Met	Lys	Pro	Pro	Ser	Ile	
				65					70					75	
Ile	Leu	Leu	Thr	Ala	Asn	Gly	Arg	Phe	Glu	Val	Gly	Lys	Lys	Ile	
				80					85					90	
Cys	Leu	Ser	Ile	Ser	Gly	His	His	Pro	Glu	Thr	Trp	Gln	Pro	Ser	
				95					100					105	
Trp	Ser	Ile	Arg	Thr	Ala	Leu	Leu	Ala	Ile	Ile	Gly	Phe	Met	Pro	
				110					115					120	
Thr	Lys	Gly	Glu	Gly	Ala	Ile	Gly	Ser	Leu	Asp	Tyr	Thr	Pro	Glu	
				125					130					135	
Glu	Arg	Arg	Ala	Leu	Ala	Lys	Lys	Ser	Gln	Asp	Phe	Cys	Cys	Glu	
				140					145					150	
Gly	Cys	Gly	Ser	Ala	Met	Lys	Asp	Val	Leu	Leu	Pro	Leu	Lys	Ser	
				155					160					165	
Gly	Ser	Asp	Ser	Ser	Gln	Ala	Asp	Gln	Glu	Ala	Lys	Glu	Leu	Ala	
				170					175					180	
Arg	Gln	Ile	Ser	Phe	Lys	Ala	Glu	Val	Asn	Ser	Ser	Gly	Lys	Thr	
				185					190					195	

PF-0502-1 DIV

Ile Ser Glu Ser Asp Leu Asn His Ser Phe Ser Leu Thr Asp Leu
200 205 210
Gln Asp Asp Ile Pro Thr Thr Phe Gln Gly Ala Thr Ala Ser Thr
215 220 225
Ser Tyr Gly Leu Gln Asn Ser Ser Ala Ala Ser Phe His Gln Pro
230 235 240
Thr Gln Pro Val Ala Lys Asn Thr Ser Met Ser Pro Arg Gln Arg
245 250 255
Arg Ala Gln Gln Gln Ser Gln Arg Arg Leu Ser Thr Ser Pro Asp
260 265 270
Val Ile Gln Gly His Gln Pro Arg Asp Asn His Thr Asp His Gly
275 280 285
Gly Ser Ala Val Leu Ile Val Ile Leu Thr Leu Ala Leu Ala Ala
290 295 300
Leu Ile Phe Arg Arg Ile Tyr Leu Ala Asn Glu Tyr Ile Phe Asp
305 310 315
Phe Glu Leu

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1006 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: PROSNOT14
- (B) CLONE: 1728211

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4 :

AGGTCTTCTC CGTACTGGAG TTTGCACCCT CAAATCATTC TTTTAAGAAA ATTGAGTTCC 60
AGCCTCCAGA AGCCAAGAAG TTCTTCAGCA CAGTGCGGAA GGAGATGGCG CTGCTGGCTA 120
CCTCACTGCC TGAGGGCATC ATGGTCAAGA CTTTTGAAGA TAGAATGGAC CTCTTCTCAG 180
CTCTCATCAA GGGCCCCACT CGAACCCCTT ACGAGGATGG CCTCTACTTG TTTGACATCC 240
AGCTCCCCAA CATCTACCCA GCCGTGCCCC CCCACTTCTG CTACCTCTCC CAATGCAGTG 300
GCCGCCTGAA CCCCACCTG TATGACAATG GGAAGGTGTG TGTCAGCCTC CTGGGCACCT 360
GGATTGGAAG GGGGACAGAG AGGTGGACAA GCAAGTCCAG CCTTCTCCAG GTGCTCATCT 420
CCATCCAAGG TCTGATCCTG GTAAATGAAC CATACTACAA CGAAGCCGGC TTCGACAGTG 480
ACCGAGGCCT GCAGGAAGGC TATGAAAACA GTCGCTGTTA CAATGAGATG GCGCTGATCC 540
GCGTGGTGCA GTCCATGACC CAGCTGGTGC GGCGGCCCCC CGAGGTCTTT GAGCAGGAGA 600
TCAGGCAACA CTTTAGCACT GGTGGCTGGC GGCTGGTGAA CCGTATCGAG TCCTGGCTGG 660
AAACCCATGC CCTGCTGGAG AAGGCCCAGG CACTGCCCAA CGGGGTGCCC AAGGCCAGCA 720
GCTCGCCAGA GCCCCAGCT GTAGCCGAGC TGTCAGACTC CGGCCAACAA GAACCTGAGG 780
ATGGAGGGCC AGCCCCAGGA GAGGCCTCCC AGGGCTCAGA CTCAGAGGGC GGTGCCCAGG 840
GCCTGGCCTT CAGCTAGCAG GGACCACACA GACCAGACTT CGGAGACCGC ACCAGACGCA 900
TCGGTGCCAC CCAGTGTGAA ACCCAAAGAA GCGGAGAAAG AGCNTTAAGA GCTACCGGAG 960
CTTCTTACCT GAGAAGAGTG GCTACCCTGA CATCGGCTTC CCCCTC 1006

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2067 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: SINTNOT13
- (B) CLONE: 1803905

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5 :

TTTNTGTGT GTGCNCCNG TNGTGNCCC TNCTANGTGG TGGGNTTTGT TGTNGGTGTC 60
 TTCCTCTGTT TNTNTTTNCC TTGCCATTGC CTCAGNCTTC CNGTCCTCNT NNNCCNNTTT 120
 TGNGCNCCCC CCNATGTCCN TGCGCCCTAG NNGTFTGGNN CNCCCCCAA GNTCTCNGNC 180
 CTNNCTGAAA GGGNNCCCC CTNNGAAAAG GCCTNCTCGG NAATTGGNGT AAAAANATCC 240
 CCCNTNGCCA AGGGNNGGNG GGGAAAACCC CGGGTTTNGG AAAGGGGGCN TTTTNCNGNG 300
 GNNATTNCCC CCCCTGTTGG GTNTGGNACC GGNNTTCCCC NGANCCTNTT TTAACNAANC 360
 TTGGGGGAAG GTTTNNCCAA NTTNTTTCGG GACTTCCCC CAACACNCNT TTTANNGGAA 420
 NGGCGNCNT ATTTCTTCGG CCGAGGTTA CATCTGATG TTGGAGCGCC GCCGCCGCA 480
 TGGCCAGCA GCAGATGACC AGCTCGAGA AGGCCCTGAT GCTCGAGCTG AAATCCCTGC 540
 AGGAGGAACC GGTGGAGGGC TTCCGGATCA CCCTGGTGA CGAGTCCGAC CTCTACAACT 600
 GGGAGGTGGC CATCTTCGGA CTCCCCAACA CCCTCTACGA AGGCGGCTAC TTCAAGGCGC 660
 ATATTAAATT TCCTATTGAC TACCCCTATT CACCACCTAC CTTCAGATTCT TTAGCAAAA 720
 TGTGGCACC CAACATTTAT GAGAAATGGAG ATGTATGCAT TTCGATTCTT CATCCGCTG 780
 TAGATGACCC ACAGAGTGA GAACTGCCTT CTGAAAGGTG GAATCCTACT CAGAATGTGA 840
 GGACTATCCT ATTAAGTGTA ATCTCACTGC TTAATGAGCC CAACACCTTC TCCCCAGCCA 900
 ATGTCGATG TTCAGTTATG TTCAGGAAAT GGAGAGACAG TAAAGGAAAA GACAAAGAAT 960
 ATGCTGAAAT TATTAGGAAA CAAGTTTCAG CCACTAAGGC CGAAGCAGAA AAGGATGGAG 1020
 TGAAGGTCCC CACAACCCTG GCGGAATACT GCATCAAAAC TAAAGTGCC TCCAATGACA 1080
 ACAGCTCAGA TTTGCTTTAC GACGACTGT ATGATGACGA CATTGATGAT GAAGATGAGG 1140
 AGGAGGAAGA TGCCGACTGT TATGATGATG ATGATTCTGG GATGAGGAGT CGTGACGTGC 1200
 TCCTTCAGTG CCCCTGTACT GCCCTGCCAT CTCAGGCCAA AGGGAGGGGA GCAAGTGGGG 1260
 ACCTGGCCAT GGCCCCTCAG CAAAAACCTA TTCACAGCGG GTGGGGAAAC ACACACAGCT 1320
 CCTGCTGACT CCCCTTATGG ATCTCAGTTT GCTCCTTTT ATGGACCTTT AATGGAGAGA 1380
 GAGTAACCT CCACAGAAATG TCTGAATTCT TGCATTCTTT ACCCTTCCAT CACTATATTG 1440
 ATTCTTTTTT TAAAAATAT GAACCCAAAC TCCCGCCTCA CTTCGTCCT ACAGAATGTT 1500
 CACAGCAAAA CACGTTTGGT CTGTTTTTAG ATTCTTGAAG AATTCAATAG TCTTTCAAGA 1560
 TGTTTAATGT GTTTAAAGCT GGAACCTGT TGGGAGTTCA CAAGTGCTGC ATATACTGGG 1620
 TAGCAAAAGA AAATGGAAAA AAACCCACAA AACAACTTT AAAAAAAAAA AAAAACAAAT 1680
 TTGCCAAGGT TTAGCTGCTC ATTTACATTA GTGTGTGTGC ATTCGTTTCAG CCCCATGGTG 1740
 GTGAATTCTG TTTCTTTCCT TTCTTAAGGC TGGGACATGG TGGGCATCAG GGACTTTGTG 1800
 CTAAGCCTGA TGAAATGTGC TCCTTCAATC TCCATGAAAC CATCGTAACA TGGAGGCCTC 1860
 AGTGTCTCTG AGGAGAGAAA TCAGACTTTG TTTTTTGAAA TCGATTGGGA TCGAAAGCCT 1920
 GAAATAAATA TTCATACTTT CCATAGTCCA CCAAAATGA GAAAGGAGGA GAAAAAAA 1980
 AAGGGGGGGG CGCCGGCCTA GTGACCCCTG TCGACCCGGG AATTAAATTC CGGACCGGGA 2040
 CCTGCAGGGG TGTACCAGGT TTTCCCT 2067

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1067 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: COLNTUT16
- (B) CLONE: 2792472

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6 :

cgaggaggcc gagccaagcc agcgaccac catggagacc cgctacaacc tgaagagtcc 60
 ggctgttaaa cgtttaatga aagaagcggc agaattgaaa gatccaacag atcattacca 120
 tgcgagcct ttagaggata acctttttga atggcacttc acggttagag ggccccaga 180
 ctccgatttt gatggaggag tttatcacgg gcggatagta ctgccaccag agtatcccat 240
 gaaaccacca agcattattc tcctaacggc taatggtcga tttgaagtgg gcaagaaaat 300
 ctgtttgagc atctcaggcc atcatctga aacttggcag ccttcgtgga gtataaggac 360
 agcattatta gccatcattg ggtttatgcc aacaaaagga gagggagcca taggttctct 420
 agattacact cctgaggaaa gaagagcact tgccaaaaaa tcacaagatt tctgttgtga 480
 aggatgtggc tctgccatga aggatgtcct gttgccttta aaatctggaa gcgattcaag 540
 ccaagctgac caagaagcca aagaactggc taggcaaata agctttaagg cagaagtcaa 600
 ttcattctga aagactatct ctgagtcaga cttaaaccac tctttttcac taactgattt 660
 acaagatgat ataccataa cattccaggc tgctacggcc agtacatgt acggactcca 720
 gaattctca gcagcatcct ttcattcaacc tacccaacct gttagctaaga atacctccat 780

PF-0502-1 DIV

gagccctcga	cagcgccggg	cccagcagca	gagtcagaga	aggttgtcta	cttcaccaga	840
tgtaatccag	ggccaccagc	caagagacaa	ccacactgat	catggtgggt	cagctgtact	900
gattgtcatc	ctgacttttg	cattggcagc	tcttatattc	cgacgaatat	atctggcaaa	960
cgaatacata	tttgactttg	agttataata	tggttttgtg	acttatgagc	tgtgactcaa	1020
ctgcttcatt	aaacattctg	cattgggtat	aatctaaaaa	aaaaaaa		1067

09930026 081401